Machine learning – An Introduction,
Center for Informatics and Computational Science (CICS),
University of Notre Dame,
Spring 2019

GAUSSIAN PROCESSES CONNECTIONS TO NEURAL NETWORK

Jiale Shi

Department of Chemical and Biomolecular Engineering, University of Notre Dame, IN 46556, U.S.A.

Email: jshi1@nd.edu, Website: `https://github.com/shijiale0609`

Gaussian processes (GPs) [1, 7, 8] are generally nonparametric, which takes an infinite number of basis functions given by a probability distribution over functions. For a non-parametric model, the required number of parameters grows with the size of the training data; therefore, GPs have a very big challenge in the computational tractability. Inference in a Gaussian process has computational complexity of $\mathcal{O}(n^3)$. Even if using low rank approximations, the computational complexity is still $\mathcal{O}(n^2 m)$, m is a user chosen parameter. A neural network (NN) [2] is a parameterised model that can be tuned via gradient descent to approximate a labelled collection of data with high precision. The computational complexity of NN is much less than GPs. However, in the process of NN, the uncertainty would be dropped in the hidden layers. The uncertainty is very important because if the the uncertainty is missing, the high precision is meaningless. In addition, overfitting is a very common problem in NN. While GPs can track the uncertainty according to the prior. Therefore, combing the advantages of two worlds, GPs and NN together can be very useful to speed up computation and track uncertainty. Recently, there are some advancements in connecting GPs and NN. [3–6]

Plan

First, I would learn the basic knowledge of both Gaussian Processes and Neural Network from books [1, 2, 7, 8]. Second, since recently there are papers [3, 4, 6] on this fields, I would read these publications and go through the mathematical derivations in these paper using the basic knowledge. Third, I would repeat the paper's results. Final, I would extend the algorithm and codes to other projects.

References

[1] C. M. Bishop. *Pattern Recognition and Machine Learning.* Springer, 2006. `http://research.microsoft.com/en-us/um/people/cmbishop/prml/`.

[2] H. B. Demuth, M. H. Beale, O. De Jess, and M. T. Hagan. *Neural network design.* Martin Hagan, 2014.

[3] M. Garnelo, D. Rosenbaum, C. J. Maddison, T. Ramalho, D. Saxton, M. Shanahan, Y. W. Teh, D. J. Rezende, and S. Eslami. Conditional neural processes. *arXiv preprint arXiv:1807.01613*, 2018.

[4] M. Garnelo, J. Schwarz, D. Rosenbaum, F. Viola, D. J. Rezende, S. Eslami, and Y. W. Teh. Neural processes. *arXiv preprint arXiv:1807.01622*, 2018.

[5] J. Hensman, N. Fusi, and N. D. Lawrence. Gaussian processes for big data. *arXiv preprint arXiv:1309.6835*, 2013.

[6] R. N. S. S. J. P. J. S.-d. Jaehoon Lee, Yasaman Bahri. Deep neural networks as gaussian processes. *International Conference on Learning Representations*, 2018.

[7] K. P. Murphy. *Machine learning: a probabilistic perspective*. MIT press, 2012.

[8] C. E. Rasmussen and C. Williams. *Gaussian Processes for Machine Learning*. MIT Press, 2006.